International Journal of Medical, Pharmacy and Drug Research (IJMPD)
Article DOI: https://dx.doi.org/10.22161/ijmpd.9.1.2
Peer-Reviewed Journal
Int. J. Med. Phar. Drug Re., 9(1), 2025
ISSN: 2456-8015

In Silico Analysis of the Promoter Regions and Regulatory Elements of DVL2, AXIN1, TCF7 & GSK3B in Triple Negative Breast Cancer

G. Deepthi Reddy[1], V. Brahmaiah[2], B. Y. Kavitha[2], G. Deepika[2]*

[1]Department of Environmental Toxicology, Institute of Genetics and Hospital for Genetic Diseases, Osmania University, MNJ Institute of Oncology & Research Cancer Centre,
[2]Department of Environmental Toxicology, Institute of Genetics and Hospital for Genetic Diseases, Osmania University, Hyderabad, Telangana, India.
*Corresponding author-deepika2104@gmail.com

Received: 29 Jan 2025; Received in revised form: 01 Mar 2025; Accepted: 10 Mar 2025; Available online: 13 Mar 2025

***Abstract*—** *Breast cancers are the leading cause of deaths in women worldwide. Breast cancer majorly occurs in female while some male, about 1% are also affected with this disease. Breast cancers are categorized into four major sub-groups based on the expression of the receptors i.e., Estrogen receptor(ER), Progesterone receptor(PR), and Human epidermal growth factor receptor(HER2) on their cell surface. Among all the sub-groups Triple Negative Breast Cancers are the most aggressive sub-type occurring in premenopausal young women usually below 40years of age. This study is focused on in silico analysis to understand the expression of the genes, it is important to understand the promoter region which plays a key role and to understand the contribution of theses genes in TNBC and the important promoter motifs. The significant promoter motifs that control the transcription of these genes are enumerated below. Studying Single Nucleotide Polymorphisms (SNPs) is critical in genetics for various reasons. SNPs are the most common type of genetic variants in human genome, where a single nucleotide base is altered in the DNA sequence. These variations can have significant impact on the human health, disease susceptibility, drug responses, and evolutionary biology. SNPs or single nucleotide polymorphisms are the changes in the DNA sequences that can have an important role in causing the disease. The SNPs are found in the promoter regions using the tool PredictSNP, and few of them thus found were deleterious. The gene variants plays important role in cancer tumorigenesis, so in this study we have screened promoter regions by using in silico tools to identify SNPs.*

Keywords*— *Triple Negative Breast Cancer, Single Nucleotide Polymorphisms, DVL2, AXIN1, TCF7, GSK3B.

I. INTRODUCTION

Breast Cancers are the largest cause of deaths in women worldwide. These cancers are categorised into two types – Ductal Carcinomas and Lobular Carcinomas. Breast cancers can be further divided into 4 sub-types based on the expression of the hormone receptors, such as estrogen(ER), progesterone(PR), human epidermal growth factor receptor(HER2). The four most widely recognized sub-types are Luminal A, Luminal B, HER2-positive, and TNBC. Of all the above mentioned sub-types TNBC is the most aggressive type occurring in the young women and in most of the women with BRCA1 gene mutations. The incidence rates of Breast cancers more precisely the percentage of TNBCs are alarming which calls for early detection techniques, diagnostic markers, and more efficient treatment therapies with less side effects and targeted therapies. The genes considered

for the study are DVL2, AXIN1, GSK3B, TCF7 which are the most important genes in understanding the pathophysiology of TNBC. All the mentioned genes are in the WNT signalling pathway. Reading the promoter regions of the genes is crucial for many reasons, various genes are having SNPs that are linked to different kinds of cancer. They effect gene expression through a variety of ways. These pathways are determined by the function of the genetic components that contain the particular SNPs.

Studying single nucleotide polymorphisms (SNPs) is critical in genetics for many reasons. SNPs are the most common type of genetic variants in human genome, where a single base is altered in the DNA sequence. These variations can have significant impact on the human health, disease susceptibility, drug responses, and evolutionary biology. Some of the SNPs in the promoter region are found using the PredictSNP tool in the DVL2 gene are rs222850, rs62059167, whereas in GSK3B rs334558, rs3755556, rs186739572, rs58853520, and in TCF7 gene rs30491, rs30492, rs43153, rs173424, rs187491297. Of these SNPs rs222850, rs30491, rs30492, rs43153, rs3755556 & rs186739572 were found to have deleterious effect while the other SNPs have a neutral effect.

II. DATA RETRIEVAL

The eukaryotic promoter sequences are retrieved from the National Council for Biotechnological Information (NCBI). The curated high-throughput experimentally validated and analyzed promoter sequences harboring promoter motifs of the genes DVL2, AXIN1, TCF7, and GSK3B are identified using *Homo sapiens* Eukaryotic Promoter Database (hsEPDnew) (https://epd.expasy.org/epd/human/human_database.php?db=human). Retrieving of the promoter sequence is important to understand the gene regulation, gene expression and it's function. It is important to study the promoter regions as they contain the binding sites for the transcription factors, since these transcription factors are responsive to the external stimuli. These regions also contain regulatory elements like enhancers, silencers, and insulators which control gene activation and repression.

IDENTIFICATION OF PROMOTER MOTIFS:

Identifying the promoter motifs is a crucial for the study as these motifs are the small stretch of the DNA sequences that are present within the promoter sequence and are the specific regions that are identified by the transcription factors and regulatory elements[2]. Promoter motifs of promoter regions recruit transcription factors and play a crucial role in regulating gene expression. TATA-box, CCAAT-box, and GC-box are the most common promoter motifs located upstream of the initiator site. TATA-box is located upstream of 25 – 30 base pairs (bp) and recognized by TATA-binding protein (TBP), a component of transcription factor II D complex. Additionally, the consensus pentanucleotide CCAAT-box located upstream of 75 – 80 bp serve as an enhancer in gene expression. They act as a binding site for CCAAT binding transcription factor (CBF)[12,21,22] that plays a crucial role in the assembly of the transcription machinery. Moreover, the housekeeping genes have GC-rich regions which serve as a binding factor for transcription factor Sp1[2,13,17].

PRIMER DESIGNING AND IN SILICO PCR:

The regulatory regions of 1200bp harboring all the experimentally validated promoter motifs of the aforementioned genes DVL2, AXIN1, TCF7, and GSK3B were considered for designing primers using Primer3[23-25]. The product size range was defined to be 851 – 1000 bp. Moreover, the primer sets were analyzed for self-dimer or cross-dimer formation using a Multiple Primer Analyzer (Thermofisher Scientific). Furthermore, the specificity of the primers was cross-checked by performing *in silico* PCR using the UCSC genome browser against the *Homo sapiens* GRCh38 genome assembly.

SINGLE NUCLEOTIDE POLYMORPHISM ANALYSIS:

SNP analysis in the present study is important to look for the association with biological traits, phenotypes, and in various diseases. SNPs are among the most common types of genetic variation and play a crucial role in understanding genetic influences on biological functions. The analysis of the SNPs become even more important when they are present in and around the promoter regions because the variations present in these regions effect the gene regulation and gene

expression. The Single Nucleotide Polymorphisms (SNPs) in the specific promoter sequence (1200 bp) with global mutation allele frequency (gMAF) > 0.02 were retrieved from dbSNP[10,11,18,19] database of NCBI. Further, the identified SNPs' nature was classified using the consensus classifier tool PredictSNP[17,18]. PredictSNP predicts the functional impact of the SNPs on the protein. It accumulates the results from multiple prediction algorithms, providing the a confidence score and the consensus sequence to evaluate whether the found SNP has a deleterious effect or a neutral effect[18].

EFFECT OF SNPs ON TRANSCRIPTION FACTOR BINDING SITE:

Gene regulation may be significantly impacted by single nucleotide polymorphisms (SNPs) in or close to transcription factor binding sites (TFBS)[1,16,18]. Even a single nucleotide alteration in these DNA patterns can affect the capacity of transcription factors (TFs) to bind, as TFs use these motifs to control gene expression, it is important to find out the SNPs using bioinformatic tools to confirm the SNPs found in the biological tissues.

The effect of SNPs on the transcription factor binding site was analyzed by using SNP2TFBS[14,15,16,18,19]. The associated consensus sequence, class, and family of identified transcription factors were validated by the JASPAR 2024 database.

III. RESULTS AND DISCUSSION

Eukaryotic Promoter Sequences:

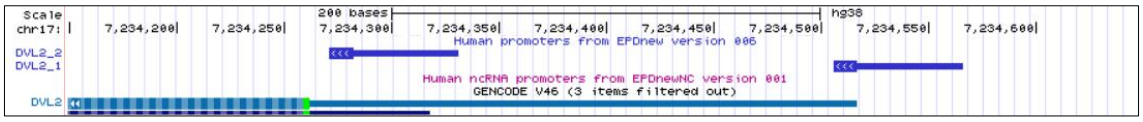
The eukaryotic promoter sequences were retrieved from the NCBI and the specific regions harboring promoter motifs were validated by the Eukaryotic Promoter Database (hsEPDnew)[20-22](Table 1). This tool gives us the comprehensive information about the experimentally validated promoters in Eukaryotic genome.It is widely used in the molecular biology, genomics, and bioinformatics for studying and understanding of the gene regulation and the structure of the promoter region. A region of 1200 bp from the reference promoter sequence harboring the promoter motifs (Figure 1a – 1d) was considered for further analysis. Additionally, we have predicted the promoter motifs (TATA-box, CCAAT-box, and GC-box) located upstream of the start site by using the hsEPDnew prediction tools (Figure 2a – 2d).

Table 1: Reference Sequence ID of Promoters and their Consensus Sequence harboring Promoter Motifs

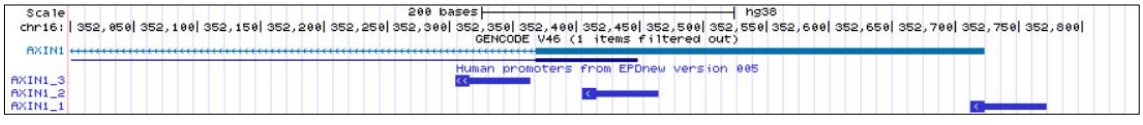
Gene Symbol	Promoter ID (region)	Consensus Sequence Harboring PMs′	No. of PMs′^{E}
DVL2	NM_004422	chr17:7233949-7235077	2
AXIN1	NM_181050	chr16:351853-352882	3
TCF7	NM_201634	chr5:134114692-134115675	2
GSK3B	NM_001146156	chr3:120093558-120094704	3

PMs′ – Promoter Motifs

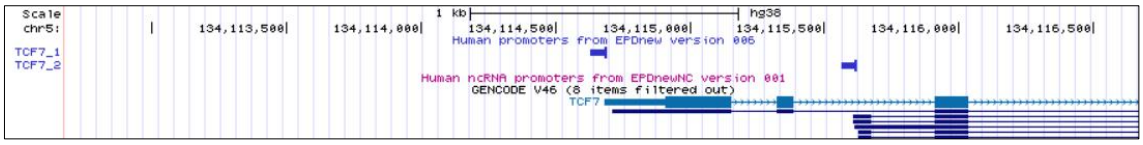
PMs′^{E}– Experimentally validated Promoter Motifs



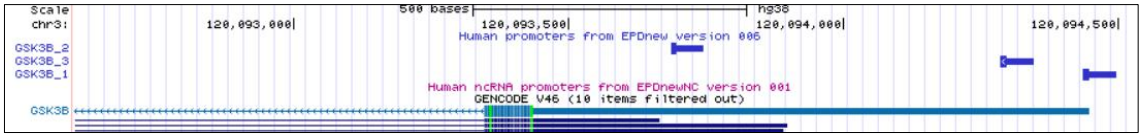
(a)DVL2



(b)AXIN1

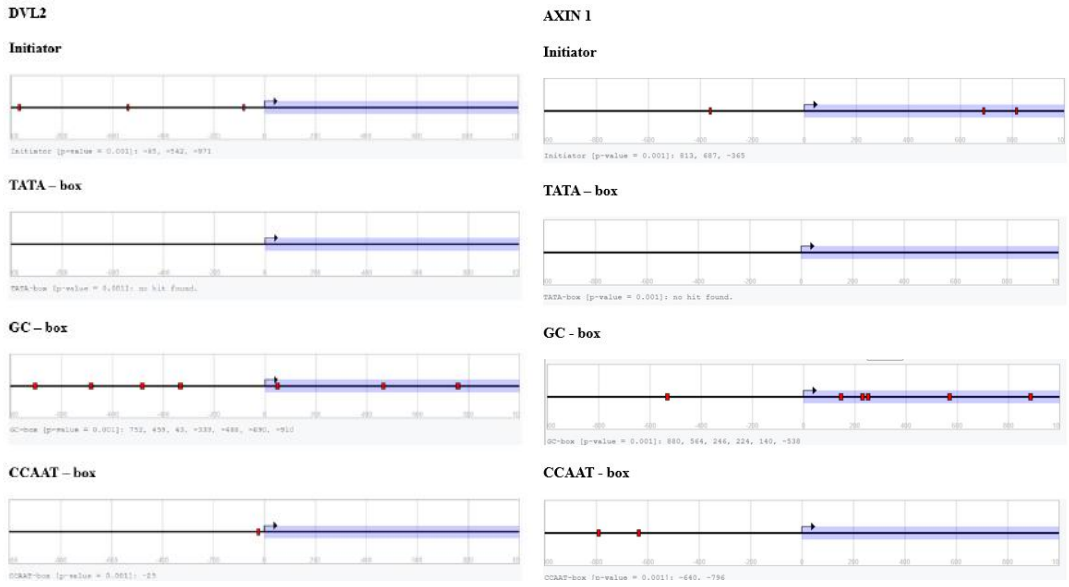


(c)TCF7

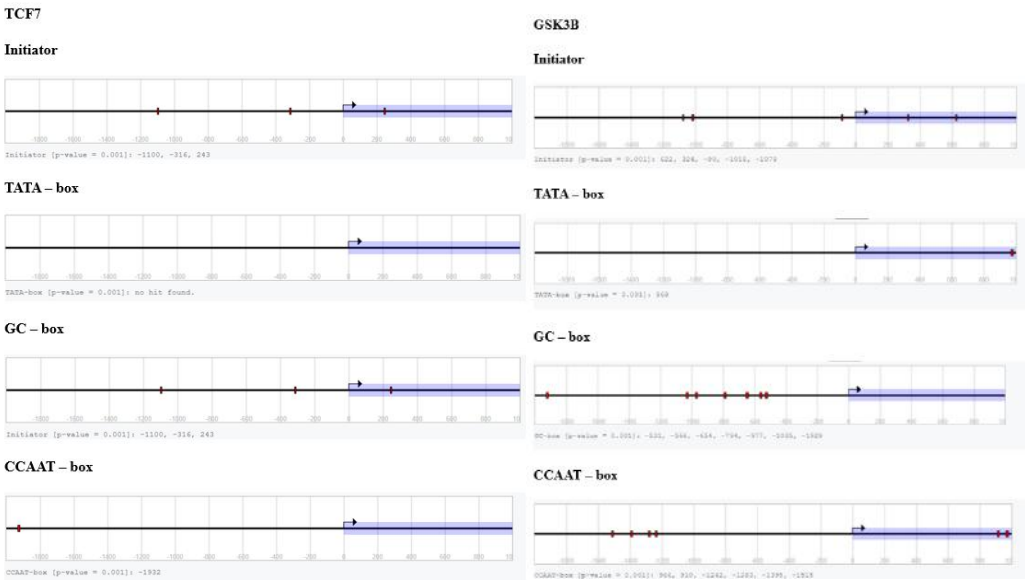


(d)GSK3B

Fig.1a – 1d: Representative structures illustrating Consensus Promoter Sequence harboring promoter Motifs.



(a) **(b)**



(c) **(d)**

Fig.2a – 2d: Representative structures illustrating the initiator site, TATA-box, GC-box and CCAAT-box in the consensus Promoter Motif sequence.

Primer Designing and In Silico PCR:

As mentioned before the region of 1200 bp from the promoter sequence was used for designing primers. The details of the primers along with the melting temperature, GC concentration, and amplicon size are given Table 2. Further, these primers are validated for self-dimer and cross-dimer formation (Table 3). The primers without any dimerization were validated for their specificity by performing in silico PCR against GRCh38 genome assembly.

Table 2: Details of the Forward and Reverse Primers along with amplicon size, melting temperature, and GC concentration.

Gene Symbol	Primers	Product Size (bp)	Tm (^{0}C)	GC%	Self / Cross Dimer formation	In Silico PCR
DVL2	LP: GACAGCCTGCGTGTTGTAAA RP: TGTCGCCCAATCCACTCTAG	1067	59.06 59.18	50 55	No	chr17:7234011-7235077 1067bp
	LP: ATTGTATTTGGCCCTCCCCA RP: CAGTGTGGCCCAAAGTAGAC	901	58.99 58.48	50 55	Yes	chr17:7233978-7234878 901bp
	LP: TTTCCCACCCCAAGACAAGT RP: AATCTGCTCCACCATAGGCC	969	59.07 59.52	50 55	No	chr17:7233949-7234917 969bp
	LP: ACCCACGTCTCAAAGTCCAA RP: ATCCACTCTAGCAAAGCCCC	943	59.17 59.45	50 55	No	chr17:7234020-7234962 943bp
	LP: TCCCCACTCGAGTCTAGGAA RP: GGCCAGAAAATCCCAGTGTG	900	59.00 59.11	55 55	Yes	chr17:7233965-7234864 900bp
AXIN1	LP: CCCCACGCTCCTCACTTTAT RP: CCACACCCAAGAGAAAGACC	966	59.46 58.10	55 55	No	chr16:351917-352882 966bp
	LP: CACACCCTTTACCTCGTCCC RP: CAGTGCAGCTCTGGTTTCAC	969	60.04 59.41	60 55	No	chr16:351853-352821 969bp
	LP: CCTCACTTTATCCTCGCGCT RP: ACCCAAGAGAAAGACCAGTGT	953	59.90 58.87	55 47.62	No	chr16:351921-352873 953bp
	LP: CACTTTATCCTCGCGCTTTCA RP: GGCTAAATTCCAAAGTGCGG	995	59.00 57.72	47.62 50.00	Yes	chr16:351876-352870 995bp
TCF7	LP: GCCTTTGATGTTCCGACCC RP: GGCGGATTAGTCAGTCACCT	954	58.82 59.18	57.89 55.00	No	chr5:134114695+134115648 954bp
	LP: GGCGCCTTTGATGTTCCG RP: TGGCCAATTCTGTCTCCTGA	981	59.51 58.64	61.11 50.00	No	chr5:134114692+134115672 981bp
	LP: CTTTGATGTTCCGACCCGCC RP: CCTTGGCCAATTCTGTCTCC	979	61.64 58.53	60.00 55.00	No	chr5:134114697+134115675 979bp
	LP: GCTCGGAGGTTCGGACTC RP: GCCAAGTTTTAGGGGAAGGAC	957	59.51 58.83	66.67 52.38	No	chr5:134114807+134115763 957bp
GSK3B	LP: TGGTTTCTCAGGCTGATCGG RP: TCCTTCAAGACAGATCGGCA	1081	59.75 58.73	55 50	Yes	chr3:120093624-120094704 1081bp
	LP: AAGGAGGTGGAGGACGAGTA RP: TCGTCCTCTTGGCTTTTCAC	961	59.30 58.12	55 50	No	chr3:120093562-120094522 961bp
	LP: AGGTGGAGGACGAGTAGGAG RP: TCCTCTTGGCTTTTCACTCCT	954	59.74 58.94	60 47.62	No	chr3:120093565-120094518 954bp
	LP: GGAGGAGCCGCAAACAAAC RP: TCTTTCCCCTCCCTTTCCTG	985	59.71 58.63	57.89 55.00	No	chr3:120093691-120094675 985bp
	LP: AGGAGTGGGGAAGTGCAAG RP: GACTTCGTCCTCTTGGCTTT	933	59.24 57.83	57.89 50	No	chr3:120093558-120094490 933bp

FP – Forward Primer; RP – Reverse Primer and Tm – Melting temperature.

Table 3: The details of the self-dimer and cross-dimer formation.

Gene Symbol	Primers	Self-dimers	Cross-dimers
DVL2	FP: ATTGTATTTGGCCCTCCCCA RP: CAGTGTGGCCCAAAGTAGAC	No	5-attgtatttggccctcccca-> | ||||| || | | <-cagatgaaacccggtgtgac-5
	FP: TCCCCACTCGAGTCTAGGAA RP: GGCCAGAAAATCCCAGTGTG	No	5-tccccactcgagtctaggaa-> | |||| | | | <-gtgtgaccctaaaagaccgg-5
AXIN1	FP: CACTTTATCCTCGCGCTTTCA RP: GGCTAAATTCCAAAGTGCGG	No	5-cactttatcctcgcgctttca-> ||| |||| <- ggcgtgaaaccttaaatcgg-5
TCF7	FP: GCTCCTCACTTTATCCTCGC RP: TAAATTCCAAAGTGCGGGGC	No	5-gctcctcactttatcctcgc-> || || ||||| | <-cggggcgtgaaaccttaaat-5
GSK3B	FP: TGGTTTCTCAGGCTGATCGG RP: TCCTTCAAGACAGATCGGCA	1 dimer for FP: 5-tggtttctcaggctgatcgg-> | |||| | <- ggctagtcggactctttggt-5	5-tggtttctcaggctgatcgg-> || |||| | <- acggctagacagaacttcct-5

SNPs in Specific Promoter Regions:

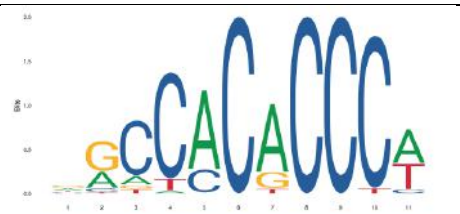
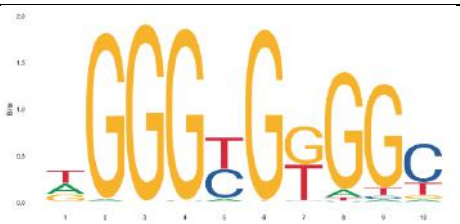
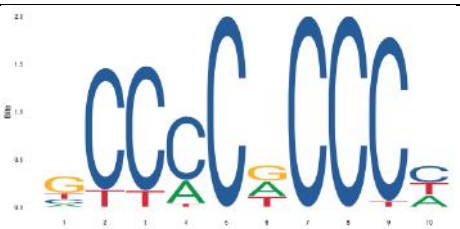
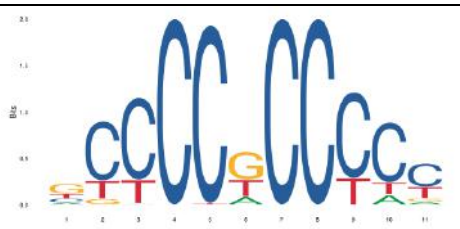
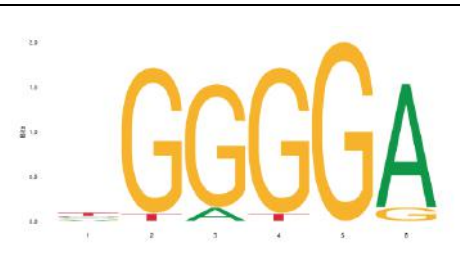
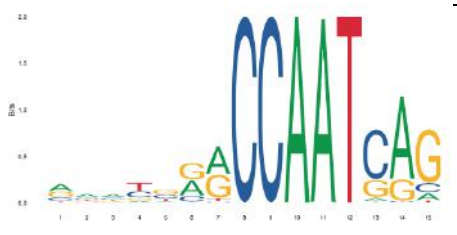
The SNPs with MAF > 0.02 in the aforementioned specific promoter regions were identified from the 1000genomes database (Table 4). Further, the PredictSNP has classified seven of them as deleterious (Table 4). In DVL2, one SNP rs62059167 of the two was found to be affecting the transcription factor binding site (TFBS) of Klf1 and Kfl4. In the case of TCF7, of the five SNPs, only one was found to be affecting TFBS of KLF5, SP1, and MZF1. Similarly, GSK3B also has one SNP rs3755556[5,4,9] that was found to be affecting NF-□B (Table 5).

Table 4: The details of Single Nucleotide Polymorphisms along with their nature in the Specific Promoter Sequences.

Gene	dbSNP ID	Location	Class	Frequency	Nature of SNP
DVL2	rs222850	17:7234631	upstream_variant	G=0.296377	Deleterious
	rs62059167	17:7234768	upstream_variant	A=0.101655	Neutral
TCF7	rs30491	5:134115522	upstream_variant	G=0.064803	Deleterious
	rs30492	5:134114887	upstream_variant	T=0.038257	Deleterious
	rs43153	5:134115448	upstream_variant	C=0.038101	Deleterious
	rs173424	5:134114806	upstream_variant	A=0.25843	Neutral
	rs187491297	5:134115098	upstream_variant	T=0.021081	Neutral
GSK3B	rs334558	3:120094435	upstream_variant	A=0.399282	Deleterious

	rs3755556	3:120094637	upstream_variant	A=0.031855	Deleterious
	rs186739572	3:120094598	upstream_variant	T=0.021081	Deleterious
	rs58853520	3:120093962	upstream_variant / Del	-=0.074016	-

Table 5: Transcription factor binding sites (Class, family and consensus sequence) effected by the SNPs (dbSNP ID).

Gene	dbSNP ID	Effected TFBS	Class	Family	Consensus Sequence
DVL2	rs62059167	Klf1	C2H2 Zinc Finger Factors	Three-zinc finger kruppel related	
		Klf4	C2H2 Zinc Finger Factors	Three-zinc finger kruppel related	
TCF7	rs30491	Klf5	C2H2 Zinc Finger Factors	Three-zinc finger kruppel related	
		SP1	C2H2 Zinc Finger Factors	Three-zinc finger kruppel related	
		MZF1	C2H2 Zinc Finger Factors	More than three-zinc finger kruppel related	
GSK3B	rs3755556	NF-YB	Heteromeric CCAAT binding factors	CCAAT binding factors	

IV. CONCLUSION

This paper provides an excellent grasp of the genetic variants present in the promoter regions of the genes DVL2, AXIN1, GSK3B & TCF7. Though these genetic variants are involved in various diseases like Alzheimers, depressive disorders and some other cancers. The overall results imply that discovered consensus sequences, frequent candidate motifs and the transcription factors by bioinformatics technique are likely predicted to provide for a better understanding of gene expression in the present study on Triple Negative Breast Cancers. Some of the SNPs in the promoter region are found using the PredictSNP tool in the DVL2 gene are rs222850, rs62059167, whereas in GSK3B rs334558, rs3755556, rs186739572, rs58853520, and in TCF7 gene rs30491, rs30492, rs43153, rs173424, rs187491297. Of these SNPs, rs222850, rs30491, rs30492, rs43153, rs334558, rs3755556 & rs186739572 were known to have deleterious effect while the other SNPs have a neutral effect.

REFERENCES

[1] Bantihun G, Kebede M. In silico analysis of promoter region and regulatory elements of mitogenome co-expressed trn gene clusters encoding for bio-pesticide in entomopathogenic fungus, Metarhizium anisopliae: strain ME1. J Genet Eng Biotechnol. 2021 Jun 22;19(1):94. doi: 10.1186/s43141-021-00191-6. PMID: 34156573; PMCID: PMC8218090.

[2] Aman Beshir J, Kebede M. In silico analysis of promoter regions and regulatory elements (motifs and CpG islands) of the genes encoding for alcohol production in Saccharomyces cerevisiaea S288C and Schizosaccharomyces pombe 972h. J Genet Eng Biotechnol. 2021 Jan 11;19(1):8. doi: 10.1186/s43141-020-00097-9. PMID: 33428031; PMCID: PMC7801573.

[3] Brittney N. Keel, William T. Oliver, John W. Keele, Amanda K. Lindholm-Perry, Evaluation of transcript assembly in multiple porcine tissues suggests optimal sequencing depth for RNA-Seq using total RNA library, Animal Gene, Volumes 17-18, 2020, 200105, ISSN 2352-4065, https://doi.org/10.1016/j.angen.2020.200105. https://www.sciencedirect.com/science/article/pii/S2352406520300051

[4] Liu S, Wang L, Sun N, Yang C, Liu Z, Li X, Cao X, Xu Y, Zhang K. The gender-specific association of rs334558 in GSK3β with major depressive disorder. Medicine (Baltimore). 2017 Jan;96(3):e5928. doi: 10.1097/MD.0000000000005928. PMID: 28099358; PMCID: PMC5279103.

[5] Lin Q, Cao YP, Gao J. Common Polymorphisms in the GSK3β Gene May Contribute to the Pathogenesis of Alzheimer Disease: A Meta-Analysis. J Geriatr Psychiatry Neurol. 2015 Jun;28(2):83-93. doi: 10.1177/0891988714554712. Epub 2014 Oct 27. PMID: 25351705.

[6] Aristizabal-Pachon AF, Castillo WO. Role of GSK3β in breast cancer susceptibility. Cancer Biomark. 2017;18(2):169-175. doi: 10.3233/CBM-160120. PMID: 27983530.

[7] Li W, Gan C, Yu S, Xu J, Tang L, Li Q, Zhu Z, Cheng H. GSK3β rs3107669 polymorphism implicates chemotherapy-associated retrospective memory deficits in breast cancer survivors. Am J Cancer Res. 2023 Oct 15;13(10):4961-4975. PMID: 37970370; PMCID: PMC10636677.

[8] Aristizábal-Pachón AF, Takahashi CS. [Effect of genetics, epigenetics and variations in the transcriptional expression of cadherin-E in breast cancer susceptibility]. Biomedica. 2016 Dec 1;36(4):593-602. Spanish. doi: 10.7705/biomedica.v36i4.3135. PMID: 27992986.

[9] Rosales-Reynoso MA, Zepeda-López P, Saucedo-Sariñana AM, Pineda-Razo TD, Barros-Núñez P, Gallegos-Arreola MP, Flores-Martínez SE, Sánchez-Corona J. GSK3β Polymorphisms Are Associated with Tumor Site and TNM Stage in Colorectal Cancer. Arch Iran Med. 2019 Aug 1;22(8):453-460. PMID: 31679348.

[10] Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams R, Trawick BW, Pruitt KD, Sherry ST. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20-D26. doi: 10.1093/nar/gkab1112. PMID: 34850941; PMCID: PMC8728269.

[11] Phan L, Zhang H, Wang Q, Villamarin R, Hefferon T, Ramanathan A, Kattman B. The evolution of dbSNP: 25 years of impact in genomic research. Nucleic Acids Res. 2025 Jan 6;53(D1):D925-D931. doi: 10.1093/nar/gkae977. PMID: 39530225; PMCID: PMC11701571.

[12] Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, Comeau DC, Funk K, Kim S, Klimke W, Marchler-Bauer A, Landrum M, Lathrop S, Lu Z, Madden TL, O'Leary N, Phan L, Rangwala SH, Schneider VA, Skripchenko Y, Wang J, Ye J, Trawick BW, Pruitt KD, Sherry ST. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2021 Jan 8;49(D1):D10-D17. doi:

10.1093/nar/gkaa892. PMID: 33095870; PMCID: PMC7778943.

[13] Hara A, Lu E, Johnstone L, Wei M, Sun S, Hallmark B, Watkins JC, Zhang HH, Yao G, Chilton FH. Identification of an allele-specific transcription factor binding interaction that regulates *PLA2G2A* gene expression. bioRxiv [Preprint]. 2023 Dec 13:2023.12.12.571290. doi: 10.1101/2023.12.12.571290. Update in: Bioinform Biol Insights. 2024 Jul 30;18:11779322241261427. doi: 10.1177/11779322241261427. PMID: 38168258; PMCID: PMC10760018.

[14] Burnham KL, Milind N, Lee W, Kwok AJ, Cano-Gamez K, Mi Y, Geoghegan CG, Zhang P; GAinS Investigators; McKechnie S, Soranzo N, Hinds CJ, Knight JC, Davenport EE. eQTLs identify regulatory networks and drivers of variation in the individual response to sepsis. Cell Genom. 2024 Jul 10;4(7):100587. doi: 10.1016/j.xgen.2024.100587. Epub 2024 Jun 18. PMID: 38897207; PMCID: PMC11293594.

[15] Kubota N, Suyama M. An integrated analysis of public genomic data unveils a possible functional mechanism of psoriasis risk via a long-range ERRFI1 enhancer. BMC Med Genomics. 2020 Jan 22;13(1):8. doi: 10.1186/s12920-020-0662-9. PMID: 31969149; PMCID: PMC6977261.

[16] Ferreira JM, Santos BRCD, Moura EL, Santos ACMD, Vencioneck Dutra JC, Figueiredo EVMS, Lima Filho JL. Narrowing the Relationship between Human CCR5 Gene Polymorphisms and Chagas Disease: Systematic Review and Meta-Analysis. Life (Basel). 2023 Aug 2;13(8):1677. doi: 10.3390/life13081677. PMID: 37629534; PMCID: PMC10455882.

[17] Fatemeh S, Mahboobeh Z, Khadijeh A, Amirhossein MK, Pegah M. An in-silico study to determine susceptibility to cancer by evaluating the coding and non-coding non-synonymous single nucleotide variants in the *SOCS3* gene. J Biomol Struct Dyn. 2024 Oct;42(16):8281-8292. doi: 10.1080/07391102.2023.2256408. Epub 2023 Sep 27. PMID: 37753777.

[18] Prabhu NB, Vinay CM, Satyamoorthy K, Rai PS. Pharmacogenomics deliberations of 2-deoxy-d-glucose in the treatment of COVID-19 disease: an in silico approach. 3 Biotech. 2022 Nov;12(11):287. doi: 10.1007/s13205-022-03363-4. Epub 2022 Sep 21. PMID: 36164436; PMCID: PMC9491670.

[19] Moraghebi M, Negahi AA, Bazireh H, Abbasi H, Ahmadi M, Sarikhani Z, Mousavi P. The Analysis of SNPs' Function in miR-21 and miR146a/b in Multiple Sclerosis and Active Lesions: An In Silico Study. Bioinform Biol Insights. 2022 Aug 4;16:11779322221116322. doi: 10.1177/11779322221116322. PMID: 35958297; PMCID: PMC9358209.

[20] Dreos R, Ambrosini G, Périer RC, Bucher P. The Eukaryotic Promoter Database: expansion of EPDnew and new promoter analysis tools. Nucleic Acids Res. 2015 Jan;43(Database issue):D92-6. doi: 10.1093/nar/gku1111. Epub 2014 Nov 6. PMID: 25378343; PMCID: PMC4383928.

[21] Dreos R, Ambrosini G, Groux R, Cavin Périer R, Bucher P. The eukaryotic promoter database in its 30th year: focus on non-vertebrate organisms. Nucleic Acids Res. 2017 Jan 4;45(D1):D51-D55. doi: 10.1093/nar/gkw1069. Epub 2016 Nov 28. PMID: 27899657; PMCID: PMC5210552.

[22] Abugessaisa I, Noguchi S, Hasegawa A, Kondo A, Kawaji H, Carninci P, Kasukawa T. refTSS: A Reference Data Set for Human and Mouse Transcription Start Sites. J Mol Biol. 2019 Jun 14;431(13):2407-2422. doi: 10.1016/j.jmb.2019.04.045. Epub 2019 May 8. PMID: 31075273.

[23] Kumar A, Chordia N. In silico PCR primer designing and validation. Methods Mol Biol. 2015;1275:143-51. doi: 10.1007/978-1-4939-2365-6_10. PMID: 25697657.

[24] Kaur S, Bishnoi R, Priyadarshini P, Singla D, Chhuneja P. DSP: database of disease susceptibility genes in plants. Funct Integr Genomics. 2023 Jun 17;23(3):204. doi: 10.1007/s10142-023-01132-x. PMID: 37329484.

[25] Tan C, Chapman B, Wang P, Zhang Q, Zhou G, Zhang XQ, Barrero RA, Bellgard MI, Li C. BarleyVarDB: a database of barley genomic variation. Database (Oxford). 2020 Nov 28;2020:baaa091. doi: 10.1093/database/baaa091. PMID: 33247932; PMCID: PMC7698660.